

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

IN RE APPLICATION OF: Koji YAMAMOTO, et al.

GAU:

SERIAL NO: New Application

EXAMINER:

FILED: Herewith

FOR: AN OBJECT DETECTION METHOD AND A VIDEO RETRIEVAL METHOD

REQUEST FOR PRIORITY

ASSISTANT COMMISSIONER FOR PATENTS  
WASHINGTON, D.C. 20231

SIR:

- ☐ Full benefit of the filing date of U.S. Application Serial Number, filed, is claimed pursuant to the provisions of 35 U.S.C. §120.
- ☐ Full benefit of the filing date of U.S. Provisional Application Serial Number, filed, is claimed pursuant to the provisions of 35 U.S.C. §119(e).
- ☒ Applicants claim any right to priority from any earlier filed applications to which they may be entitled pursuant to the provisions of 35 U.S.C. §119, as noted below.

In the matter of the above-identified application for patent, notice is hereby given that the applicants claim as priority:

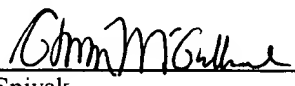
| <u>COUNTRY</u> | <u>APPLICATION NUMBER</u> | <u>MONTH/DAY/YEAR</u> |
|----------------|---------------------------|-----------------------|
| Japan          | 11-022372                 | January 29, 1999      |

Certified copies of the corresponding Convention Application(s)

- ☒ are submitted herewith
- ☐ will be submitted prior to payment of the Final Fee
- ☐ were filed in prior application Serial No. filed
- ☐ were submitted to the International Bureau in PCT Application Number .  
Receipt of the certified copies by the International Bureau in a timely manner under PCT Rule 17.1(a) has been acknowledged as evidenced by the attached PCT/IB/304.
- ☐ (A) Application Serial No.(s) were filed in prior application Serial No. filed ; and  
(B) Application Serial No.(s)
  - ☐ are submitted herewith
  - ☐ will be submitted prior to payment of the Final Fee

Respectfully Submitted,

OBLON, SPIVAK, McCLELLAND,  
MAIER & NEUSTADT, P.C.

  
Marvin J. Spivak  
Registration No. 24,913  
C. Irvin McClelland  
Registration Number 21,124

Fourth Floor  
1755 Jefferson Davis Highway  
Arlington, Virginia 22202  
Tel. (703) 413-3000  
Fax. (703) 413-2220  
(OSMMN 11/98)



13423732



日 本 国 特 許 庁

PATENT OFFICE  
JAPANESE GOVERNMENT

JC675 U.S. PTO  
09/493013



別紙添付の書類に記載されている事項は下記の出願書類に記載されて  
いる事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed  
with this Office.

出 願 年 月 日

Date of Application:

1999年 1月29日

出 願 番 号

Application Number:

平成11年特許願第022372号

出 願 人

Applicant (s):

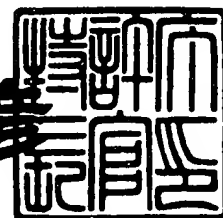
株式会社東芝

CERTIFIED COPY OF  
PRIORITY DOCUMENT

1999年11月19日

特許庁長官  
Commissioner,  
Patent Office

近 藤 隆 彦



出証番号 出証特平11-3081263

【書類名】 特許願

【整理番号】 A009900192

【提出日】 平成11年 1月29日

【あて先】 特許庁長官 殿

【国際特許分類】 G06F 15/00

【発明の名称】 映像情報記述方法、映像検索方法及び映像検索装置

【請求項の数】 17

【発明者】

    【住所又は居所】 神奈川県川崎市幸区小向東芝町 1 番地 株式会社東芝研  
    究開発センター内

    【氏名】 山本 晃司

【発明者】

    【住所又は居所】 神奈川県川崎市幸区小向東芝町 1 番地 株式会社東芝研  
    究開発センター内

    【氏名】 堀 修

【発明者】

    【住所又は居所】 神奈川県川崎市幸区小向東芝町 1 番地 株式会社東芝研  
    究開発センター内

    【氏名】 金子 敏充

【発明者】

    【住所又は居所】 神奈川県川崎市幸区小向東芝町 1 番地 株式会社東芝研  
    究開発センター内

    【氏名】 三田 雄志

【特許出願人】

    【識別番号】 000003078

    【氏名又は名称】 株式会社 東芝

【代理人】

    【識別番号】 100058479

    【弁理士】

【氏名又は名称】 鈴江 武彦

【電話番号】 03-3502-3181

【選任した代理人】

【識別番号】 100084618

【弁理士】

【氏名又は名称】 村松 貞男

【選任した代理人】

【識別番号】 100068814

【弁理士】

【氏名又は名称】 坪井 淳

【選任した代理人】

【識別番号】 100092196

【弁理士】

【氏名又は名称】 橋本 良郎

【選任した代理人】

【識別番号】 100091351

【弁理士】

【氏名又は名称】 河野 哲

【選任した代理人】

【識別番号】 100088683

【弁理士】

【氏名又は名称】 中村 誠

【選任した代理人】

【識別番号】 100070437

【弁理士】

【氏名又は名称】 河井 将次

【手数料の表示】

【予納台帳番号】 011567

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 映像情報記述方法、映像検索方法及び映像検索装置

【特許請求の範囲】

【請求項 1】

映像中の特定のオブジェクトに関する特徴量及び該映像中の背景に関する特徴量を映像情報として記述することを特徴とする映像情報記述方法。

【請求項 2】

映像中の特定のオブジェクトに関する特徴量と該映像中の背景に関する特徴量及びその差分を映像情報として記述することを特徴とする映像情報記述方法。

【請求項 3】

前記オブジェクトに関する特徴量として少なくとも前記オブジェクトの位置、形状及び動きの情報を記述し、前記背景に関する特徴量として少なくとも該背景の動きの情報を記述することを特徴とする請求項 1 及び 2 のいずれか 1 項に記載の映像情報記述方法。

【請求項 4】

請求項 1 乃至 3 のいずれか 1 項に記載の映像情報記述方法により記述された前記オブジェクトに関する特徴量及び背景に関する特徴量を映像データと共に、あるいは該映像データとは別に格納した記録媒体。

【請求項 5】

入力された映像の動ベクトルを抽出する動ベクトル抽出ステップと、  
前記動ベクトルを用いて前記映像の背景の動きを推定する推定ステップと、  
前記推定した背景の動きを除去して前記映像中の特定のオブジェクトに関する動ベクトルを検出し、該オブジェクトの領域を検出する検出ステップと、  
を有することを特徴とするオブジェクト検出方法。

【請求項 6】

前記推定ステップは、前記背景の動きを所定の変換モデルで近似し、該変換モデルの変換係数を前記映像の動ベクトルから推定することによって前記背景の動きを推定することを特徴とする請求項 5 記載のオブジェクト検出方法。

【請求項 7】

前記推定ステップは、前記変換モデルの変換係数をロバスト推定法により推定することを特徴とする請求項 6 記載のオブジェクト検出方法。

【請求項 8】

前記推定ステップは、前記映像の画面内の動ベクトルを類似度に従って領域分割し、分割された各領域を該動ベクトルの類似度に基づいてクラスタリングし、最大のクラスタの領域を背景の領域と判定する処理を含むことを特徴とする請求項 5 乃至 7 のいずれか 1 項記載のオブジェクト検出方法。

【請求項 9】

前記推定ステップは、前記映像の複数フレームの動ベクトルを各フレーム内の類似度に従って領域分割し、これらの領域をフレーム間で対応付け、対応付けられた各領域が同じクラスタに属するように各領域を該動ベクトルに基づいて各フレーム内でクラスタリングし、最大のクラスタの領域を背景の領域と判定する処理を含むことを特徴とする請求項 5 乃至 7 のいずれか 1 項記載のオブジェクト検出方法。

【請求項 10】

検索対象となる映像中の特定のオブジェクトに関する特徴量及び該映像中の背景に関する特徴量を記述しておき、

前記映像中のオブジェクトに関する特徴量から前記背景に関する特徴量を差し引いた後、外部より入力されたオブジェクトに関する特徴量と比較することによって、前記検索対象となる映像から前記外部より入力されたオブジェクトと同一のオブジェクトもしくは前記外部より入力されたオブジェクトと同一のオブジェクトを含む前記映像中のフレームの少なくとも一方を検索することを特徴とする映像検索方法。

【請求項 11】

検索対象となる映像中の特定のオブジェクトに関する特徴量と該映像中の背景に関する特徴量及びそれぞれの特徴量の差分を記述しておき、

前記差分と外部より入力されたオブジェクトに関する特徴量とを比較することによって、前記検索対象となる映像から前記外部より入力されたオブジェクトと同一のオブジェクトもしくは前記外部より入力されたオブジェクトと同一のオブ

ジェクトを含む前記映像中のフレームの少なくとも一方を検索することを特徴とする映像検索方法。

【請求項 12】

検索対象となる映像中の特定のオブジェクトに関する特徴量及び該映像中の背景に関する特徴量を記述しておき、

前記背景に関する特徴量と外部より入力された映像中の背景に関する特徴量とを比較することによって、前記検索対象となる映像から前記外部より入力された映像を得たときのカメラワークとほぼ同一のカメラワークが用いられているフレームを検索することを特徴とする映像検索方法。

【請求項 13】

検索対象となる映像中の特定のオブジェクトに関する少なくとも動きの情報を含む特徴量及び該映像中の背景に関する特徴量を記述しておき、

前記映像中の連続する複数フレームにおけるオブジェクトの動き情報と外部より入力されたオブジェクトの一連の動きの情報とを比較することにより、前記検索対象となる映像から前記外部より入力されたオブジェクトと同一のオブジェクトもしくは前記外部より入力されたオブジェクトと同一のオブジェクトを含むフレームの少なくとも一方を検索することを特徴とする映像検索方法。

【請求項 14】

検索対象となる映像中の特定のオブジェクトに関する特徴量及び該映像中の背景に関する特徴量を記述しておき、

前記映像中のオブジェクトに関する特徴量から前記背景に関する特徴量を差し引いた後、外部より入力されたオブジェクトに関する特徴量と比較することによって、前記検索対象となる映像から前記外部より入力されたオブジェクトと同一のオブジェクトもしくは前記外部より入力されたオブジェクトと同一のオブジェクトを含む前記映像中のフレームの少なくとも一方を検索することを特徴とする映像検索装置。

【請求項 15】

検索対象となる映像中の特定のオブジェクトに関する特徴量と該映像中の背景に関する特徴量及びそれぞれの特徴量の差分を記述しておき、



前記差分と外部より入力されたオブジェクトに関する特徴量とを比較することによって、前記検索対象となる映像から前記外部より入力されたオブジェクトと同一のオブジェクトもしくは前記外部より入力されたオブジェクトと同一のオブジェクトを含む前記映像中のフレームの少なくとも一方を検索することを特徴とする映像検索装置。

【請求項 1 6】

検索対象となる映像中の特定のオブジェクトに関する特徴量及び該映像中の背景に関する特徴量を記述しておき、

前記背景に関する特徴量と外部より入力された映像中の背景に関する特徴量とを比較することによって、前記検索対象となる映像から前記外部より入力された映像を得たときのカメラワークとほぼ同一のカメラワークが用いられているフレームを検索することを特徴とする映像検索装置。

【請求項 1 7】

検索対象となる映像中の特定のオブジェクトに関する少なくとも動きの情報を含む特徴量及び該映像中の背景に関する特徴量を記述しておき、

前記映像中の連続する複数フレームにおけるオブジェクトの動き情報と外部より入力されたオブジェクトの一連の動きの情報とを比較することにより、前記検索対象となる映像から前記外部より入力されたオブジェクトと同一のオブジェクトもしくは前記外部より入力されたオブジェクトと同一のオブジェクトを含むフレームの少なくとも一方を検索することを特徴とする映像検索装置。

【発明の詳細な説明】

【0 0 0 1】

【発明の属する技術分野】

本発明は、映像中のオブジェクトに注目した映像情報の記述方法及びこれを用いて特定のオブジェクトやそのオブジェクトを含むフレームの検索を行う映像検索方法及び映像検索装置に関する。

【0 0 0 2】

【従来の技術】

ディジタル衛星放送やケーブルテレビの普及に伴う放送の多チャンネル化によ

り、ユーザの入手できる映像情報は増加の一途を辿っている。一方で、計算機技術の進歩やDVDに代表される大容量記録媒体の実用化により、大量の映像情報をデジタル情報として蓄積し、計算機上で取り扱うことが容易になりつつある。

#### 【0003】

ユーザが実際に映像情報を利用するために、このように大量の映像情報の中から目的の映像に対して効率的にアクセスを行うには、有効な映像検索技術が必要となる。そのような映像検索技術として、映像中のオブジェクトになんらかの情報を付随させ、ユーザの必要とする情報を満たすオブジェクトを含む映像を検索してユーザの視聴に供する方法が考えられている。映像中のオブジェクトに情報を付随させるためには、映像からオブジェクトを抽出する処理が必要となる。しかし、増加する一方の映像情報に対して手作業でオブジェクトの抽出を行うことは、現実的ではない。

#### 【0004】

オブジェクトの自動検出技術に関しては、例えば文献（米山、中島、柳原、菅野「MPEGビデオストリームからの移動物体の検出」信学論Vol. J81-D-II, No. 8, pp. 1776-1786, 1998-08）に、背景の静止した映像からオブジェクトを検出する方法が提案されている。しかし、この方法は背景が静止していることを前提としており、背景が動く場合にはオブジェクトの検出を行うことは難しい。

#### 【0005】

すなわち、オブジェクトの形状が予め与えられていたとしても、背景の動きが分からなければ、オブジェクトの動きを用いて検索を行う際に、カメラワークの影響を受けることにより正確な動きを用いた検索ができない。例えば、左に動いているオブジェクトを追跡するように撮影した場合、オブジェクトは画面内ではほぼ静止し、背景が相対的に右に動く。そのため、画面内で左に動くオブジェクトを含む映像を検索することはできないことになる。

#### 【0006】

【発明が解決しようとする課題】

上述したように、従来の映像検索技術では背景が動いているオブジェクトを検出することができず、そのようなオブジェクトを含む映像を検索することができないという問題点があった。

【0007】

本発明は、背景が動いているオブジェクトを含む映像についても検索を可能とするための映像情報記述方法を提供することを目的とする。

【0008】

本発明の他の目的は、背景が動いているオブジェクトを検出できるオブジェクト検出方法を提供することにある。

【0009】

本発明のもう一つの目的は、背景が動いているオブジェクトを含む映像に対する種々の検索を可能とした映像検索方法及び映像検索装置を提供することにある。

【0010】

【課題を解決するための手段】

上記の課題を解決するため、本発明に係る映像情報記述方法は、映像中の特定のオブジェクトに関する特徴量及び該映像中の背景に関する特徴量を映像情報として記述することを基本的な特徴としている。

【0011】

また、本発明に係る他の映像情報記述方法は、映像中の特定のオブジェクトに関する特徴量と該映像中の背景に関する特徴量に加え、さらに両者の特徴量の差分を映像情報として記述することを特徴とする。また、映像中の特定のオブジェクトに関する特徴量と該映像中の背景に関する特徴量との差分及び背景に関する特徴量を映像情報として記述するようにしてもよい。

【0012】

ここで、特定のオブジェクトに関する特徴量としては少なくとも該オブジェクトの位置、形状及び動きの情報を記述し、また背景に関する特徴量としては少なくとも背景の動きの情報を記述することが望ましい。

【0013】

また、本発明によれば、このようにして記述されたオブジェクトに関する特徴量及び背景に関する特徴量、さらには両者の特徴量の差分を映像データと共に、あるいは映像データとは別に格納した記録媒体が提供される。

## 【 0 0 1 4 】

本発明に係るオブジェクト検出方法は、入力された映像の動ベクトルを抽出する動きベクトル抽出ステップと、この抽出した動ベクトルを用いて映像の背景の動きを推定する推定ステップと、この推定した背景の動きを除去して映像中の特定のオブジェクトに関する動ベクトルを抽出し、該オブジェクトの領域を検出する検出ステップとを有することを特徴とする。

## 【 0 0 1 5 】

ここで、背景の動きを推定する推定ステップは、背景の動きを所定の変換モデル（例えばアフィン変換や透視変換など）で近似し、その変換モデルの変換係数を映像の動ベクトルから推定することによって背景の動きを推定することを特徴とする。変換モデルの変換係数は、例えばロバスト推定法により推定される。

## 【 0 0 1 6 】

背景の動きを推定する推定ステップにおいては、映像の画面内の動ベクトルを類似度に従って領域分割し、分割された各領域を該動ベクトルの類似度に基づいてクラスタリングし、最大のクラスタの領域を背景の領域と判定する処理を含んでもよい。

## 【 0 0 1 7 】

また、背景の動きを推定する推定ステップにおいては、映像の複数フレームの動ベクトルを各フレーム内の類似度に従って領域分割し、これらの領域をフレーム間で対応付け、対応付けられた各領域が同じクラスタに属するように各領域を該動ベクトルに基づいて各フレーム内でクラスタリングし、最大のクラスタの領域を背景の領域と判定する処理を含んでもよい。

## 【 0 0 1 8 】

本発明によると、映像中のオブジェクトに関する特徴量及び背景に関する特徴量、さらには両者の特徴量の差分を記述することにより、背景が動いているオブジェクトを含む映像に対する種々の検索を行うことが可能である。

## 【0019】

第1に、検索対象となる映像中の特定のオブジェクトに関する特徴量及び該映像中の背景に関する特徴量を記述しておき、検索対象となる映像中のオブジェクトに関する特徴量から背景に関する特徴量を差し引いた後、外部より入力されたオブジェクトに関する特徴量と比較することによって、検索対象となる映像から外部より入力されたオブジェクトと同一のオブジェクトもしくは外部より入力されたオブジェクトと同一のオブジェクトを含む映像中のフレームの少なくとも一方を検索することができる。

## 【0020】

第2に、検索対象となる映像中の特定のオブジェクトに関する特徴量と該映像中の背景に関する特徴量及びそれぞれの特徴量の差分を記述しておき、この差分と外部より入力されたオブジェクトに関する特徴量とを比較することによって、検索対象となる映像から外部より入力されたオブジェクトと同一のオブジェクトもしくは外部より入力されたオブジェクトと同一のオブジェクトを含む映像中のフレームの少なくとも一方を検索することができる。

## 【0021】

第3に、検索対象となる映像中の特定のオブジェクトに関する特徴量及び該映像中の背景に関する特徴量を記述しておき、背景に関する特徴量と外部より入力された映像中の背景に関する特徴量とを比較することによって、検索対象となる映像から外部より入力された映像を得たときのカメラワークとほぼ同一のカメラワークが用いられているフレームを検索することができる。

第4に、検索対象となる映像中の特定のオブジェクトに関する少なくとも動きの情報を含む特徴量及び該映像中の背景に関する特徴量を記述しておき、映像中の連続する複数フレームにおけるオブジェクトの動き情報と外部より入力されたオブジェクトの一連の動きの情報とを比較することにより、検索対象となる映像から外部より入力されたオブジェクトと同一のオブジェクトもしくは外部より入力されたオブジェクトと同一のオブジェクトを含むフレームの少なくとも一方を検索することができる。

## 【0022】

【発明の実施の形態】

以下、図面を参照して本発明の実施の形態を説明する。

【第1の実施形態】

本実施形態では、大別して3つの機能を提供している。第1に、映像データを再生する機能の他に、映像中で動きを伴うオブジェクトを自動的に検出し、楕円や矩形などの図形を重ねて合成表示することによって、ユーザにその存在を知らせる機能を提供する。

【0023】

第2に、検出されたオブジェクトの位置、大きさ、動きなどの特徴量と背景に関する特徴量を分離して外部ファイルなどに表記データとして記述する機能を提供する。

【0024】

第3に、検出されたオブジェクトに関する特徴量のデータあるいは予め外部ファイルなどに記述された特徴量の表記データと、検索対象として外部から与えられた検索対象オブジェクトの特徴量のデータを比較し、該当するオブジェクトをユーザに提示することによって、映像中のオブジェクトを検索する機能を提供する。

【0025】

図1に、本実施形態に係る映像検索システムの構成及び処理の手順をフローチャートとして示す。

【0026】

まず、DVDなどの媒体から再生された元映像データ100を入力し（ステップ101）、この元映像データ101から後に詳しく説明する方法によって映像中の特定のオブジェクトを検出する（ステップ102）。この際、後述するように映像中の背景に関する情報も併せて検出する。検出されたオブジェクトをこれを囲むように生成された楕円や矩形などの図形と合成し、オブジェクト検出結果表示データ104として出力する（ステップ103）。

【0027】

一方、ステップ102で検出されたオブジェクトの位置、形状（大きさを含む

）及び動きなどのオブジェクトに関する特徴量を示す特徴量データと、背景の動きなどの背景に関する特徴量を示す特徴量データを記述するための特徴量データ生成処理を行い（ステップ105）、さらに生成されたオブジェクト及び背景に関する特徴量データ107を外部に出力して表記データとして記述する処理を行う（ステップ106）。

【0028】

なお、ステップ105、106においてはオブジェクトに関する特徴量データと背景に関する特徴量データを記述してもよいが、さらに両者の特徴の差分のデータを生成して記述してもよく、また場合によってはこの差分のデータと背景に関する特徴量のデータ、あるいは差分のデータとオブジェクトに関する特徴量データを生成して記述してもよい。

【0029】

ステップ106の記述処理とは、具体的には特徴量データ107を各種記録媒体やメモリに格納（記録）したり、表示したりすることをいう。特徴量データ107を格納する記録媒体は、元映像データ100を格納したDVDなどの媒体であってもよく、これとは別の記録媒体でも構わない。

【0030】

次に、オブジェクトの検索を行うために、ステップ105で生成されたオブジェクトに関する特徴量データと、ステップ109により入力された検索対象特徴量データ110との類似度判定を行い（ステップ108）、さらに類似度判定結果をオブジェクト検索結果表示データ112として合成表示するための合成表示処理を行う（ステップ111）。検索対象特徴量データ110は、検索しようとするオブジェクトの位置、形状（大きさを含む）及び動きなどの特徴量を表すデータである。

【0031】

なお、これら一連の処理は、ソフトウェアまたはハードウェアのいずれでも実現が可能である。

【0032】

次に、図2を用いて図1のステップ102のオブジェクト検出処理について詳

細に説明する。

まず、入力された元映像データ 100 より動ベクトルの抽出を行う（ステップ 201）。元映像データ 100 が M P E G 圧縮データの場合は、P ピクチャより得られる動ベクトルを用いる。この場合、動ベクトルはマクロブロック毎に与えられる。元映像データ 100 がアナログデータや動ベクトルを持たないデジタルデータの場合は、必要に応じてデジタル化し、オプティカルフローを用いて動ベクトルを抽出したり、M P E G 圧縮データに変換してから動ベクトルを抽出する。

#### 【0033】

こうして抽出された動ベクトルは、必ずしも実際のオブジェクトの動きを反映したものとはなっていない場合があり、画面の周辺部やテクスチャが平坦な部分でそれが顕著である。そこで、信頼性の低い動ベクトルを除去する処理を行う（ステップ 202）。この処理は、以下のように行われる。

#### 【0034】

まず、画面の周辺部に関しては、予め領域を定めておき、その領域に含まれる動ベクトルを除去する。一方、平坦なテクスチャを持つ部分に関しては、元映像データ 100 が M P E G 圧縮データの場合は図 3 に示すように I ピクチャの D C T（離散コサイン変換）係数の D C 成分を用い、図 4 に示すように 1 つのマクロブロックに含まれる 4 つの D C 成分の分散が閾値以下のマクロブロックの集まりを低信頼性領域として、始点はその領域のマクロブロックに含まれる動ベクトルを信頼性の低い動ベクトルとして除去する。

#### 【0035】

このようにして得られた動ベクトルのデータは、オブジェクトの動きにカメラワークなどによる背景の動きが含まれているので、正確なオブジェクトの動きを得るためには、背景の動きを除去する必要がある。そこで、本実施形態ではカメラワークなどによる背景の動きを近似する変換モデルとしてアフィン変換モデルを用い、動ベクトルを用いてその変換係数を推定することによって、背景の動きを推定する処理を行う（ステップ 203）。背景の動きのアフィン変換係数を推定する処理については、数種類の手法が存在するが、それらについては後述する



## 【0036】

次に、推定されたアフィン変換係数を用いて各動ベクトルの始点を変換し、その移動分を元の動ベクトルより差し引くことによって、背景の動きを除去する処理を行う（ステップ204）。

## 【0037】

このようにして得られた、背景の動きを含まない動ベクトルデータを類似した動ベクトルより構成される領域に分割する処理を行う（ステップ205）。具体的には、隣接する2つの動ベクトルの余弦（方向）と大きさを比較し、その差が予め定めた閾値以下であれば、同じ領域に分割する処理を全ての隣接する動ベクトルの組み合わせについて行う。

## 【0038】

このようにして得られた領域には、オブジェクトとして扱うには相応しくない小領域が含まれるので、閾値処理によりこれらを取り除く判定処理306を行い、最終的なオブジェクトデータ307を出力する。

## 【0039】

以下、動ベクトルから背景の動きを近似するステップ203のアフィン変換係数推定処理の3つの手法について説明する。

## 【0040】

## &lt;手法1&gt;

手法1では、信頼性の低い動ベクトルを除いた画面内の全ての動ベクトルを用いてアフィン変換係数を推定する。 $i$ 番目のマクロブロックの中心を $y_i$ とし、そのマクロブロックに対応する動ベクトルを $v_i$ とする。このとき、ベクトルの始点 $x_i = y_i - v_i$ のアフィン変換変形モデルによる移動先は、アフィン変換係数を $\alpha$ とすると $r_i = x_i \alpha$ となり、実際の移動先である $y_i$ との誤差は $e_i = r_i - y_i$ となる。推定残差の総和は次式のようになり、これを最小とする $\alpha$ を求めればよい。

## 【0041】

【数 1】

$$\sum_i \Psi(e_i / \sigma_i) = \min$$

【0 0 4 2】

このような問題を解くための手法として最小二乗法があり、その場合には式 (1) で  $\Psi(z) = z^2$  を用いればよい。しかし、最小二乗法を用いた場合、背景の動ベクトルとオブジェクトの動ベクトルを同列に扱うため、背景の動ベクトルのみからアフィン変換係数を推定することができず、オブジェクトの動きを含んだアフィン変換係数となってしまう。

【0 0 4 3】

そこで、画面内における背景領域の占める面積が 50% 以上あると仮定し、オブジェクトの動ベクトルを外乱とみなして、背景部分の動ベクトルからのみアフィン変換係数を推定する。外乱に強い手法として、文献：中川徹・小柳義夫著「最小二乗法による実験データ解析」東京大学出版会、に開示されているようなロバスト推定法を用いる。ここでは特に、ロバスト推定法の一つである Biweight 法による M 推定を用いる。Biweight 法では、誤差の大きい要素の重みを下げることによって、外乱の影響を受け難くしている。具体的には、先の式 (1) の  $\Psi(z)$  において、重み  $w$  を用いた次式 (2) を使用する。定数  $c$  は 5 ~ 9 に選ぶのがよいとされている。

【0 0 4 4】

【数 2】

$$\Psi(z) = \int w_j z \, dz$$

$$w_j = \begin{cases} (1 - (z_j / c)^2)^2 & |z_j| < c \\ 0 & \text{otherwise} \end{cases}$$

## 【0045】

## &lt;手法2&gt;

背景の動きを近似するアフィン変換係数推定処理において、手法2を用いた場合の処理手順を図5に示すフローチャートを用いて説明する。

## 【0046】

まず、信頼性の低いベクトルを除去する処理を行った後の動ベクトルデータ500に対して、先の手法1と同様な処理を用いて、隣接する動ベクトルが類似した領域に分割する処理を行う（ステップ501）。ただし、手法1と異なり、この時点で背景の動きを除去する処理は行われていない。

## 【0047】

次に、分割された各領域に含まれる動ベクトルより、領域の動きをアフィン変換モデルで近似したときのアフィン変換係数の推定処理を行う（ステップ502）。このときの推定処理には、手法1と同様なロバスト推定法を用いる。

## 【0048】

次に、分割された各領域のクラスタリング処理を行う（ステップ503）。これには全ての領域の組み合わせよりなるテーブルを用意して、各領域間の距離をアフィン変換係数より求める。ここでは、アフィン変換モデルの6係数のユークリッド距離を用いるが、他の距離を用いてもよい。次に、この距離が最も小さい2つの領域を統合し、統合された領域に対して新たなアフィン変換係数を求め、テーブルから統合された2領域を削除し、統合された領域を追加して、テーブルを更新する。この処理を領域間の距離が予め定めた閾値より大きくなるか、領域が1つになるまで繰り返し行う。

## 【0049】

このようにしてクラスタリングされた領域のうち、クラスタが最大の領域を背景の領域と判定する処理を行い（ステップ504）、その領域のアフィン変換係数を背景の動きのアフィン変換係数505として出力する。

## 【0050】

## &lt;手法3&gt;

背景の動きを近似するアフィン変換係数推定処理において、手法3を用いた場

合の処理手順を図 6 に示すフローチャートを用いて説明する。

【 0 0 5 1 】

まず、複数のフレームを一度に読み込み、各フレームに対して手法 2 と同様な処理を用いて隣接する動ベクトルが類似するような領域に分割する処理を行う（ステップ 6 0 1）。

【 0 0 5 2 】

次に、各領域の動きをアフィン変換モデルで近似したときの変換係数の推定処理を行い（ステップ 6 0 2）、さらに領域の位置や動ベクトルデータ及び変換係数に基づいて、フレーム間に対応する領域を求める処理を行った後（ステップ 6 0 3）、手法 2 と同様なクラスタリング処理により各フレーム内の領域をクラスタリングする（ステップ 6 0 4）。

【 0 0 5 3 】

フレーム間対応付け処理で対応付けられた領域が別のクラスタにクラスタリングされた場合は、最も多くのクラスタにクラスタリングされた結果を正解とし、他のクラスタにクラスタリングされた領域を移動する補正処理を行う（ステップ 6 0 5）。

【 0 0 5 4 】

最後に、複数フレーム間で最も面積が大きかった領域を背景とする判定処理を行い（ステップ 6 0 6）、各フレームにおける背景領域の変換係数 5 0 7 を求める。この手法 3 は、背景領域が特定のフレームで一時的に他の領域より小さくなった場合でも正しく推定ができるという利点がある。

【 0 0 5 5 】

上述の例では、背景の動きを推定する処理に用いる変換モデルにアフィン変換を用いたが、透視変換など、他の変換モデルを使用してもよい。

【 0 0 5 6 】

次に、図 7 を用いて図 1 のステップ 1 0 6 におけるオブジェクト及び背景に関する特徴量データの記述処理で用いられるデータ表現について説明する。ここでは、図 7（a）に示されるように例として第 1 0 0 0 番目のフレーム内の映像 7 0 5 に含まれる 3 つのオブジェクトの表記データ 7 0 0 を表している。この表記

データ 7 0 0 は、元映像データの映像ストリーム 7 0 6 内の対応するフレームを示すフレーム情報 7 0 1 とオブジェクトに関する特徴量 7 0 3 及び背景に関する特徴量 7 0 4 の各データより構成され、次表記データへのポインタ 7 0 2 を用いたリスト構造により管理されている。

## 【 0 0 5 7 】

オブジェクトに関する特徴量 7 0 3 は、少なくともオブジェクトの位置、形状（大きさを含む）及び動きの情報を含んでおり、具体的には例えば図 7（b）に例示されるような様々な特徴量より構成される。この例では、オブジェクトに関する特徴量 7 0 3 は「位置」、形状である「概形」、動きの情報である「アフィン変換係数」及び「動ベクトルの平均、方向」、さらに「色ヒストグラム」などから構成される。

## 【 0 0 5 8 】

ここで、オブジェクトの概形は、楕円や矩形などの簡易な図形により近似してもよい。アフィン変換係数は、前述したようにオブジェクトの動きをアフィン変換モデルで近似したときに推定される係数である。動ベクトルの平均は、オブジェクト中の動ベクトルの大きさの平均値である。また、オブジェクトの色情報が取得可能である場合には、オブジェクトの領域の色ヒストグラムが特徴量として用いることができる。オブジェクトの動きに関しては、背景の動きを除去したもの、除去しないもののいずれを記録してもよい。

## 【 0 0 5 9 】

この例のようにオブジェクトが複数存在する場合は、各オブジェクト特徴量 7 0 3 に個別の ID 番号を付け、拡張容易な例えば図 7（a）に示されるようなリスト構造により管理することが望ましい。このようなリスト構造を用いると、オブジェクト特徴量の追加や削除が容易である。

## 【 0 0 6 0 】

背景に関する特徴量 7 0 4 に関してもオブジェクトに関する特徴量 7 0 3 と同様に図 7（c）に例示されるような様々な特徴量、例えば「アフィン変換係数」、「動ベクトルの平均、方向」、「カメラワーク種」、「色ヒストグラム」などにより構成される。カメラワーク種とは、パンやズームなどの撮影に用いられる

典型的なカメラワークの種類をいう。

【0061】

次に、図8に示すフローチャートを用いて図1のステップ108における類似度判定処理について説明する。

この類似度判定処理は、元映像データに含まれる各オブジェクトに関する特徴量データ800に対して、順次、外部より入力された特徴量データ804と比較することにより行う。外部より入力する特徴量データ804は、数値などで直接データとして与えてもよいし、映像から特徴量を抽出して特徴量データとして与えてもよい。

【0062】

オブジェクトが複数の種類の特徴量を持つ場合は、各特徴量について順次、類似度判定処理により類似度を求める（ステップ803）。

【0063】

元映像データ800に含まれる特徴量データ800と外部より入力される特徴量データ804の比較は、特徴量の種類に基づいて適当な手法を用いる。例えば、特徴量が色のヒストグラムであれば、ヒストグラムの各要素の差を用いることなどが考えられる。比較するオブジェクトが異なる種類の特徴量を持つ場合は、一致する特徴量のみについて比較すればよい。

【0064】

ステップ801、802で全てのオブジェクトについて全ての特徴量データについて検索が終わったと判定されると、該当オブジェクトの情報について検索結果表示処理を行い（ステップ805）、処理は終了する。

【0065】

オブジェクトの動きを比較するときには、背景に関する特徴量データを用いて、背景の動きを除去して比較することも可能である。図9を用いて、背景の動きを分離した検索の効果について説明する。

【0066】

図9に示すように、元映像データ901は左に移動するオブジェクトを追跡するようにカメラを移動しながら撮影したものであるが、映像中では見かけ上オブ

ジェクトが静止し、背景が右に移動しているように見える。左に移動するオブジェクトを検索するために、外部よりオブジェクト 9 0 5 のデータが入力された場合、映像データ 9 0 1 のオブジェクトは静止しているために特徴量が一致せず、検索することができない。

## 【 0 0 6 7 】

しかし、本発明に従いオブジェクトに関する特徴量と背景に関する特徴量を記述すると、背景の動きを利用して、カメラワークにより動く背景 9 0 4 を分離する処理 9 0 2 により、オブジェクト本来の左方向への動きを伴うオブジェクト 9 0 3 を検出することができる。すなわち、処理 9 0 2 においてはオブジェクトに関する特徴量と背景に関する特徴量との差分を求めることにより、オブジェクト 9 0 3 のみを検出する。

## 【 0 0 6 8 】

従って、この検出したオブジェクト 9 0 2 と外部より入力されたオブジェクト 9 0 5 を比較することによって、外部より入力されたオブジェクト 9 0 5 と同一のオブジェクトを入力映像データ 9 0 1 から検索したり、外部より入力されたオブジェクト 9 0 5 と同一のオブジェクトを含む映像フレームを元映像データ 9 0 1 から検索したりすることが可能となる。この場合、前述のように差分のデータとを記述しておけば、処理 9 0 2 は不要となる。

## 【 0 0 6 9 】

また、本発明に従いオブジェクトに関する特徴量と背景に関する特徴量を記述すると、図 1 0 に示すように外部より入力されるカメラワークに一致したカメラワークの映像を検索することができる。すなわち、図 1 0 に示すように元映像データ 1 0 0 1 からオブジェクト 1 0 0 3 を分離する処理 1 0 0 2 により、カメラワークにより動く背景 1 0 0 4 のみを検出する。そして、この検出した背景 1 0 0 4 と外部より入力されるカメラワークにより動く背景 1 0 0 5 と比較することによって、外部より入力されるカメラワークと同一のカメラワークが用いられている映像フレームを元映像データ 1 0 0 1 から検索する。

## 【 0 0 7 0 】

この場合も、前述のように差分のデータとを記述しておけば、処理 1 0 0 2 は

不要となる。

【0071】

〔第2の実施形態〕

次に、図11に示すフローチャートを参照して本発明の第2の実施形態について説明する。

本実施形態では、先の第1の実施形態におけるオブジェクトの検出及び記述に代えて、予め解析された特徴量データが付加された元映像データ1100を入力し（ステップ1101）、この元映像データ1100からオブジェクトに関する特徴量データを分離して抽出する（ステップ1102）。

【0072】

以下、第1の実施形態と同様に、ステップ1102で抽出された元映像データ1100に関する特徴量データと、ステップ1109により入力された検索対象特徴量データ1110との類似度判定処理を行い（ステップ1108）、その結果をオブジェクト検索結果表示データ1112として合成表示するための合成表示処理を行う（ステップ1111）。

【0073】

なお、これら一連の処理は、ソフトウェアまたはハードウェアのいずれでも実現が可能である。

【0074】

〔第3の実施形態〕

次に、図12に示すフローチャートを用いて本発明の第3の実施形態について説明する。

本実施形態では、外部より入力された一連の動きと複数フレームに渡る表記データを比較し、時系列的な動きによるオブジェクトの検索を可能にするために、複数の連続する表記データ1201に含まれるオブジェクトのうち、同一のオブジェクトを対応付ける処理を行う（ステップ1202）。一方、外部より入力された動きデータ1203から、表記データ1201と同一の間隔で動きデータを抽出するサンプリング処理を行う（ステップ1204）。

【0075】



そして、互いに対応する表記データとサンプリングされた外部入力の動きデータを比較し（ステップ 1105）、合致するオブジェクトを含む映像を検索結果として表示する（ステップ 1106）。

【0076】

図 13 を用いて、図 12 のステップ 1202 の連続する表記データ 1201 に含まれるオブジェクトの対応付け処理について説明する。

第 N 番目の表記データに含まれるオブジェクト 1301 に関する特徴量（位置及び動き）を用いて、第 N+1 番目の表記データにおけるオブジェクトの予測される位置 1302 を求める。そして、この予測される位置 1302 に最も近い位置に存在する第 N+1 番目の表記データに含まれるオブジェクト 1303 をオブジェクト 1301 に対応するオブジェクトとする。

【0077】

図 14 を用いて、図 12 のステップ 1204 における外部から入力された動きデータ 1203 のサンプリング処理について説明する。

外部より入力された動きデータ 1401（1203 と同じ）は連続的な動きデータであるため、そのままでは数フレーム毎に付加される離散的なデータである表記データ中と比較することはできない。そこで、動きデータ 1401 を表記データのフレーム間隔でサンプリングして、このサンプリングされた動きデータ 1402 と表記データを比較する。

【0078】

なお、これら一連の処理は、ソフトウェアまたはハードウェアのいずれでも実現が可能である。

【0079】

【発明の効果】

以上説明したように、本発明によればオブジェクトに関する特徴量と背景に関する特徴量を記述しておくことにより、背景の動きを除去してオブジェクト本来の動きによる検索が可能である。

【0080】

また、大量に蓄積された映像データに対して人手に頼らずに自動的にオブジェ

クトを検出してその特徴量を抽出し、外部から入力された別の特徴量に合致したオブジェクトを検索したり、外部から入力されたカメラワークによる動きを伴う背景と同じ背景の動きを含むフレームを検索するなど、個々のユーザの目的に合った映像検索を容易に行うことができる。

【0081】

さらに、予め検出した特徴量を記述しておくことにより、検索の度に特徴量抽出処理を行う必要がなく、高速な検索が可能であり、またユーザ側にオブジェクト検出機能がない場合でも上述した検索が可能となる。

【図面の簡単な説明】

【図1】 本発明の第1の実施形態に係る映像検索システムの基本的な処理手順を示すフローチャート

【図2】 同実施形態におけるオブジェクト検出の処理手順を示すフローチャート

【図3】 同実施形態におけるオブジェクト検出に用いるMPEGストリームのIピクチャ及びPピクチャの関係を示す図

【図4】 同実施形態におけるオブジェクト検出での低信頼性ペクトルの除去について説明するための図

【図5】 同実施形態における背景領域の変換係数を求める手法を説明するためのフローチャート

【図6】 同実施形態における背景領域の変換係数を求める他の手法を説明するためのフローチャート

【図7】 同実施形態におけるオブジェクト記述処理で用いる特徴量データの構造を示す図

【図8】 同実施形態におけるオブジェクト検索の処理手順を示すフローチャート

【図9】 同実施形態におけるオブジェクト検索処理におけるカメラワークの除去について示す図

【図10】 同実施形態におけるオブジェクト検索処理における入力されたカメラワークと同一カメラワークが用いられているフレームの検索について示す

図

【図 1 1】 本発明の第 2 の実施形態に係る映像検索システムの基本的な処理手順を示すフローチャート

【図 1 2】 本発明の第 3 の実施形態に係る映像検索システムの基本的な処理手順を示すフローチャート

【図 1 3】 同実施形態における連続する表記データ間でのオブジェクトの対応付けを説明するための図

【図 1 4】 同実施形態における外部より入力された動きデータのサンプリング方法を説明するための図

【符号の説明】

7 0 0 …表記データ

7 0 3 …オブジェクトに関する特徴量

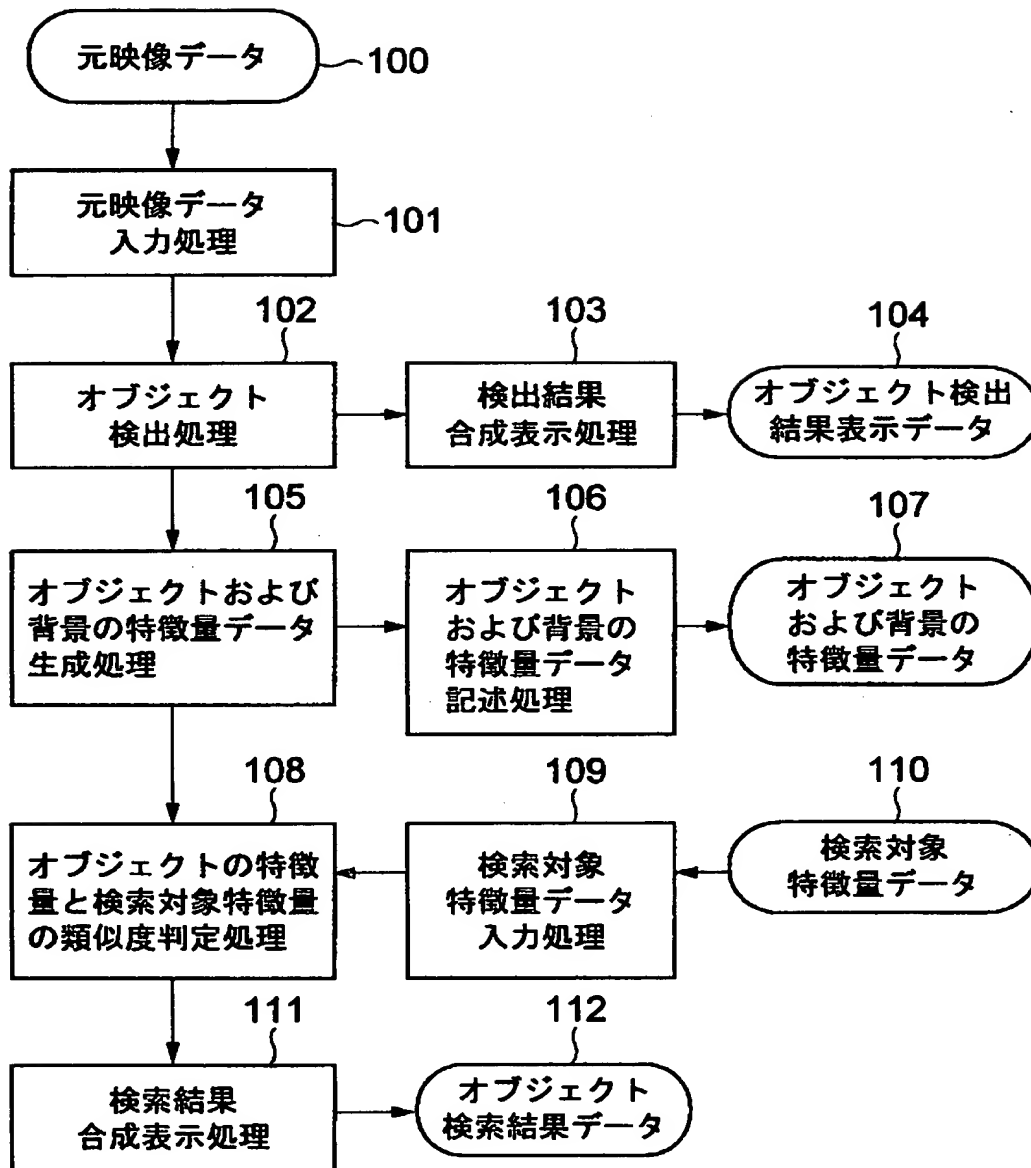
7 0 4 …背景に関する特徴量

7 0 6 …映像ストリーム

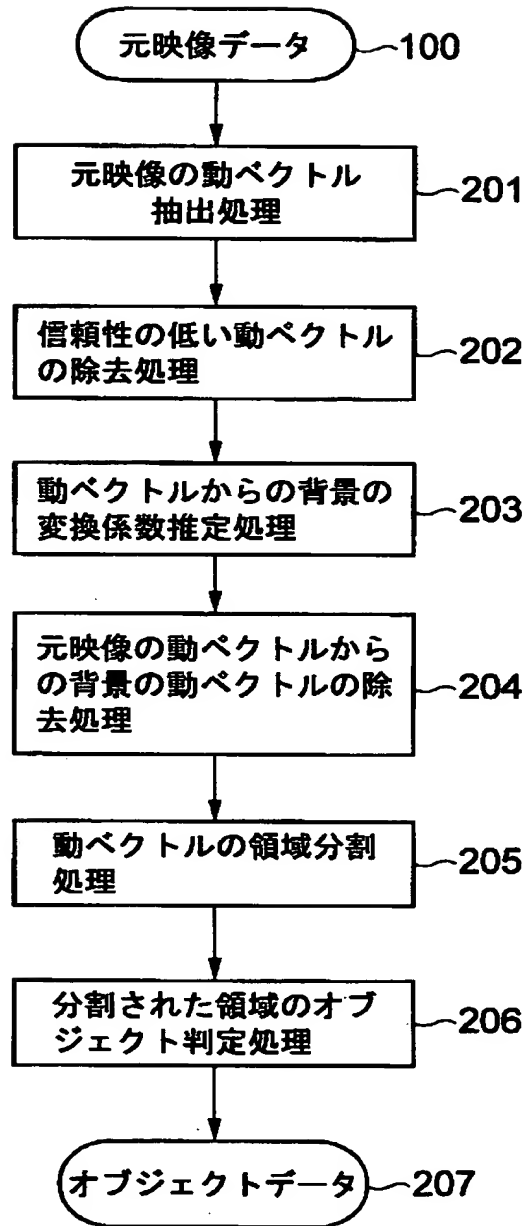
【書類名】

図面

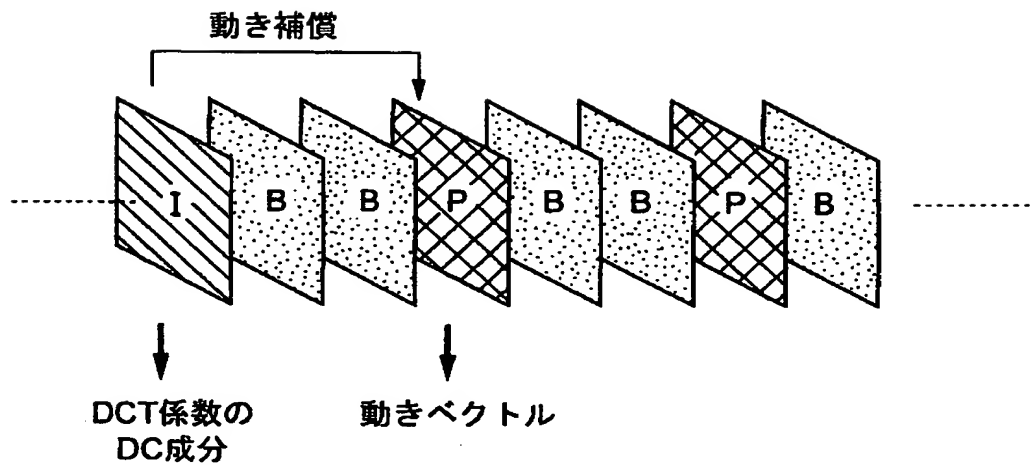
【図 1】



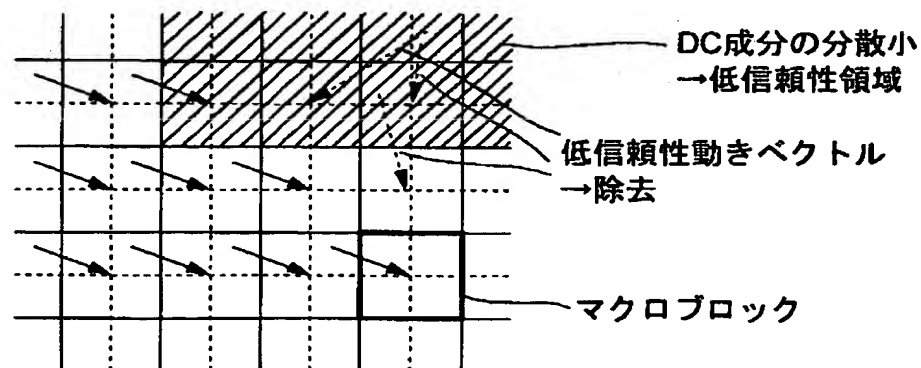
【図 2】



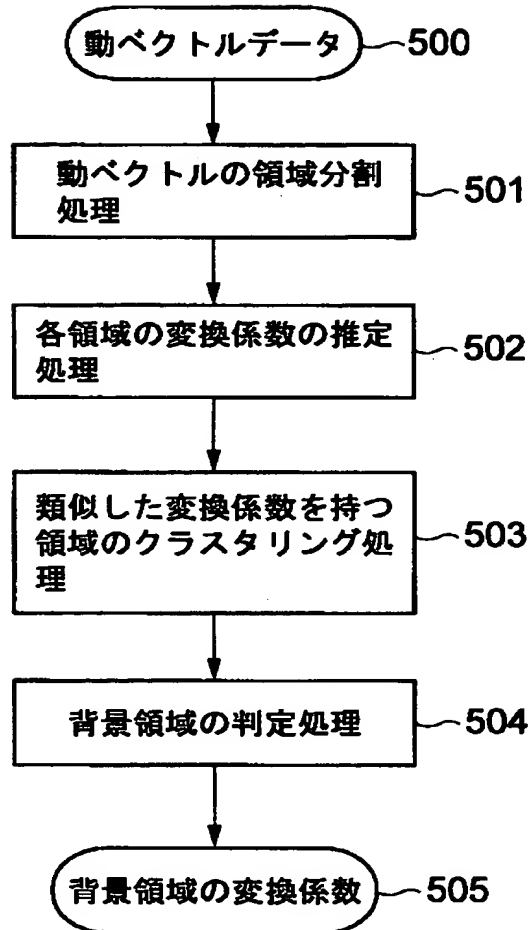
【図 3】



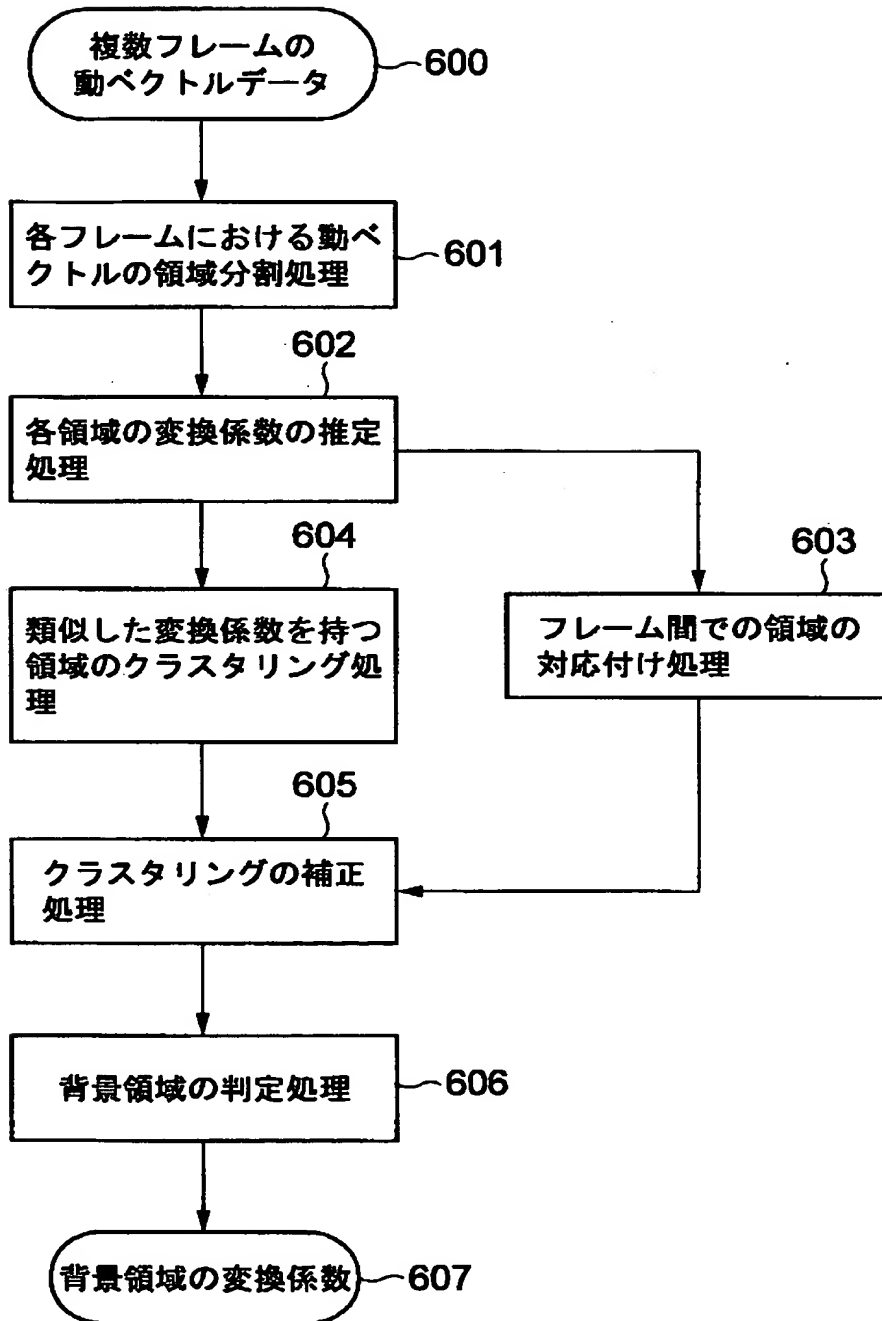
【図 4】



【図 5】

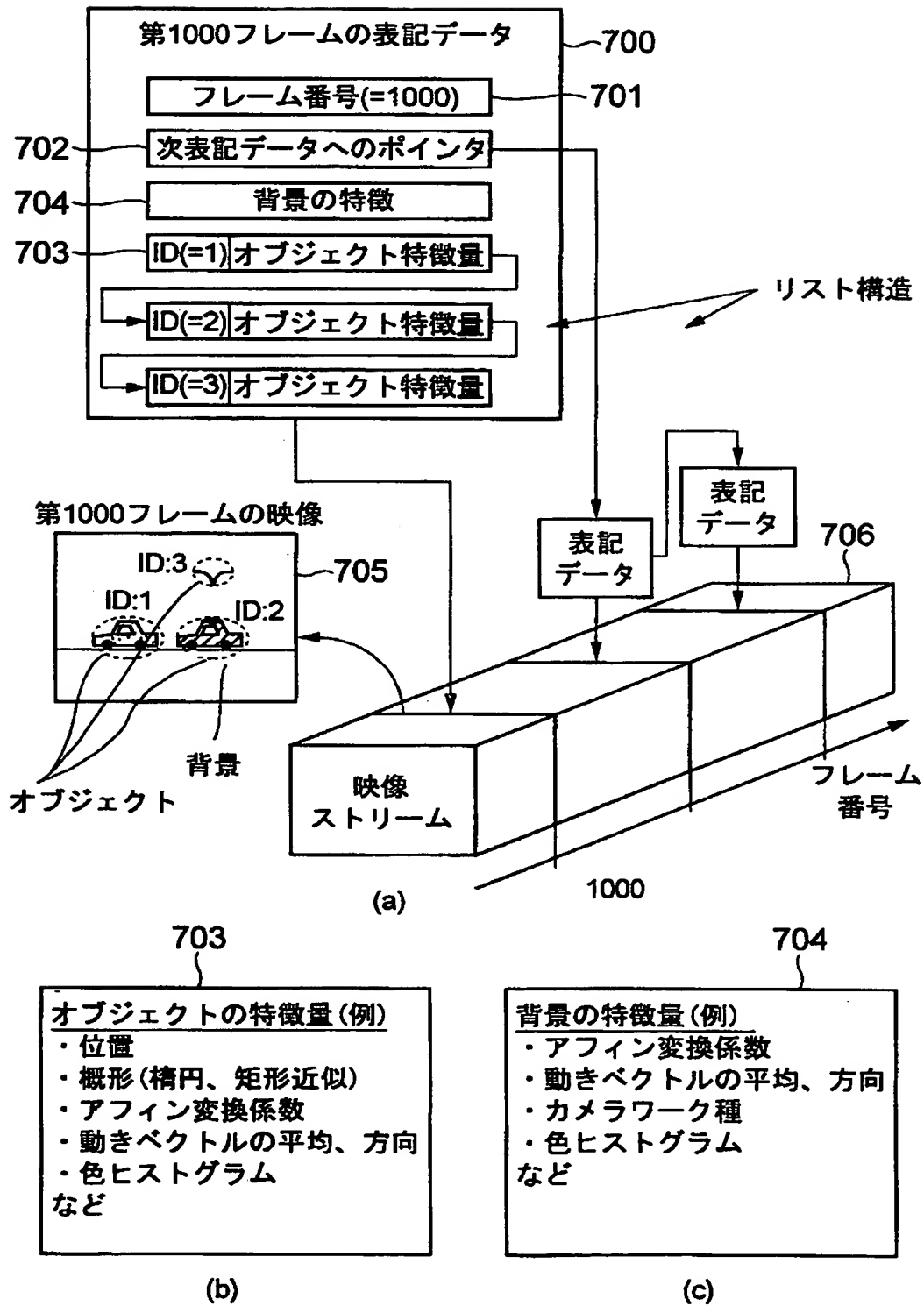


【図 6】

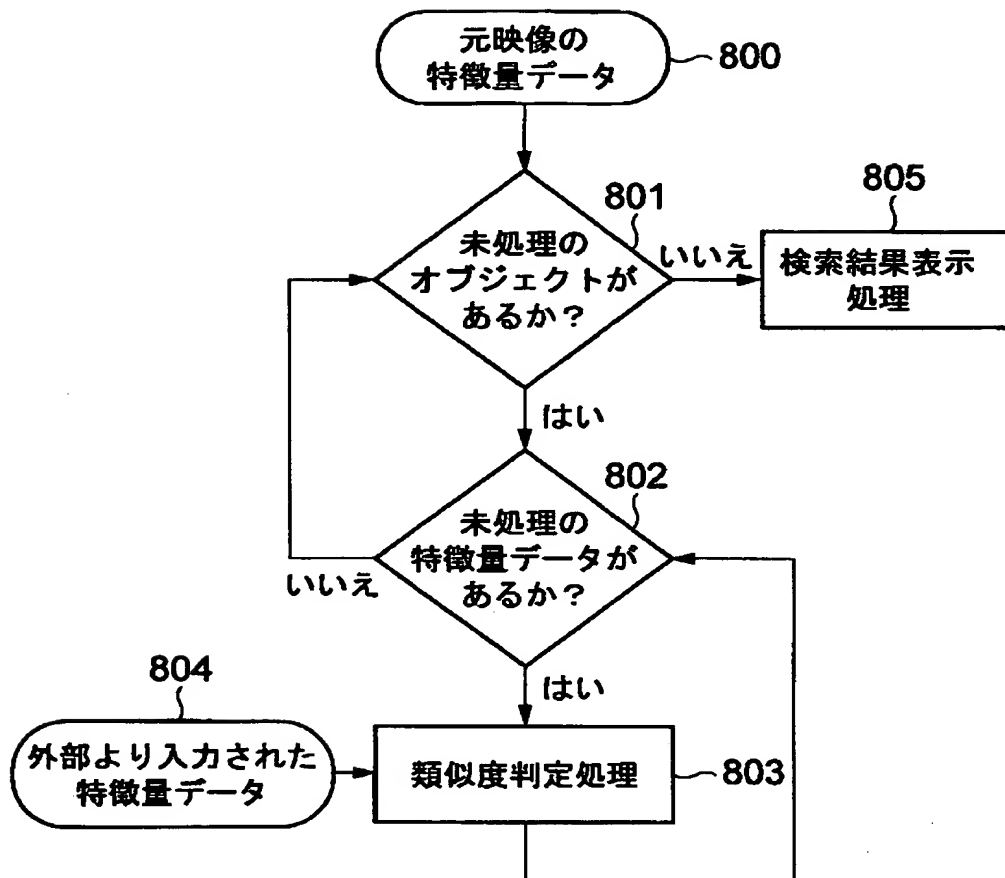




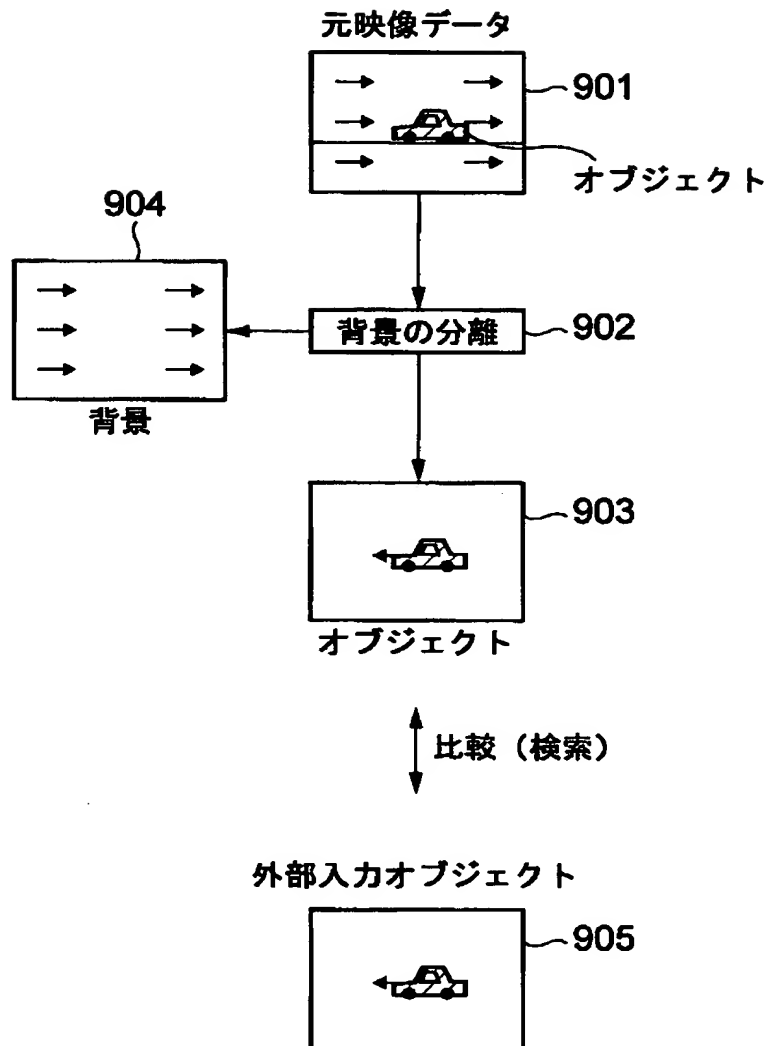
【図 7】



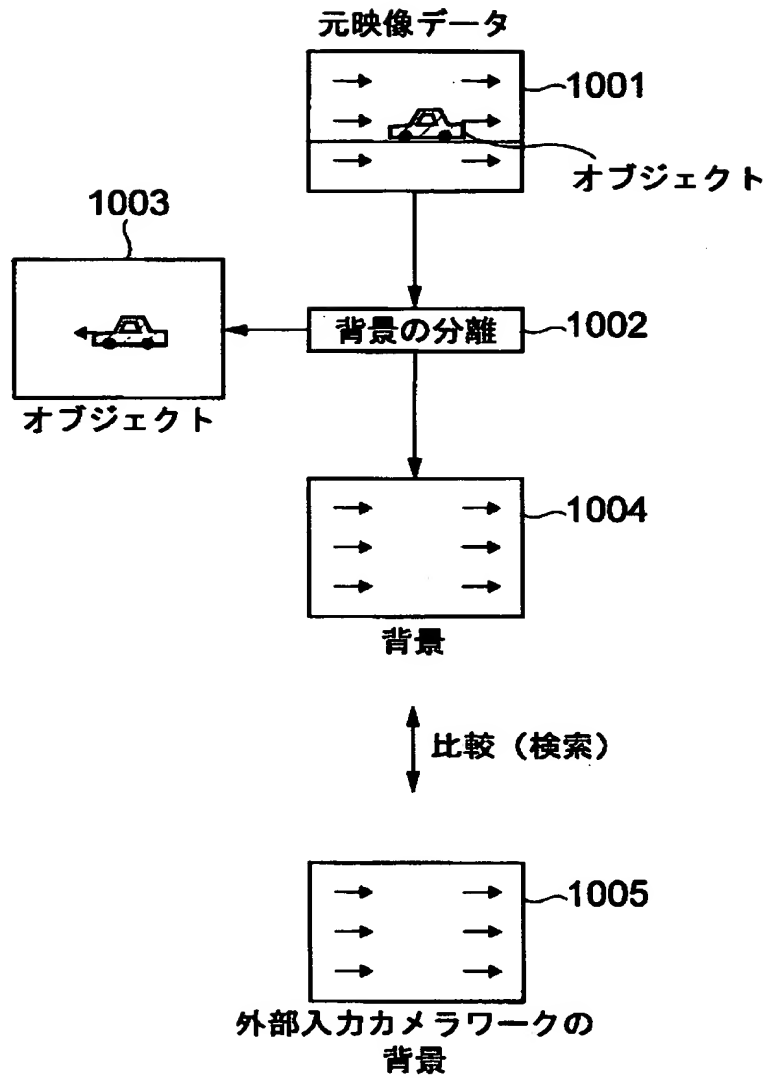
【図 8】



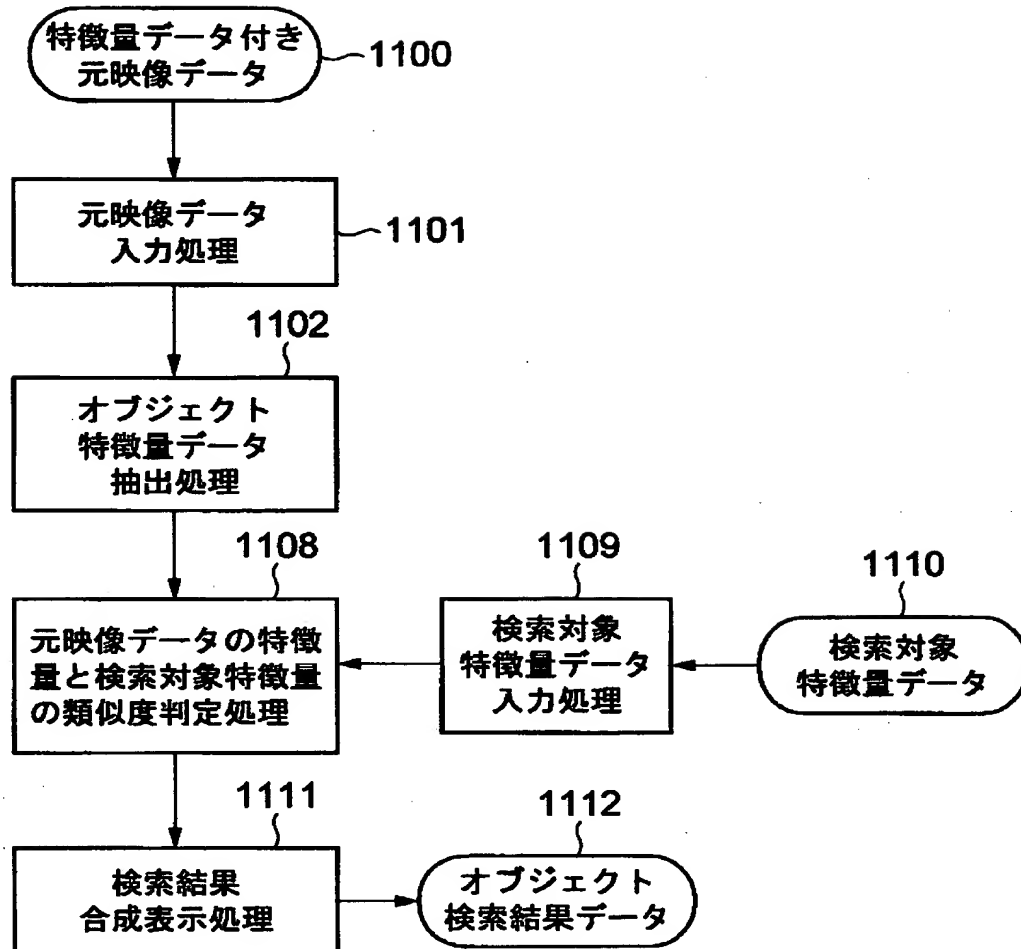
【図 9】



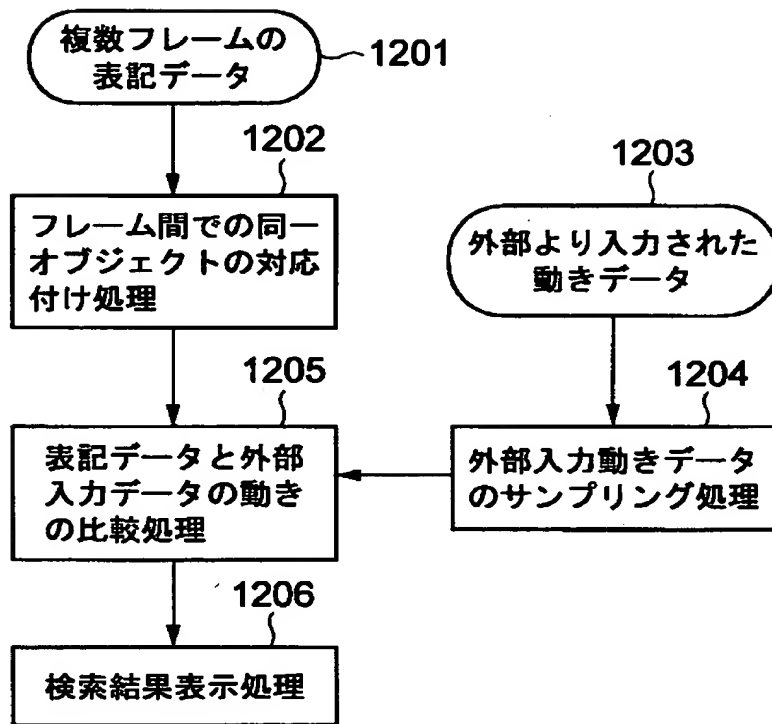
【図 1 0】



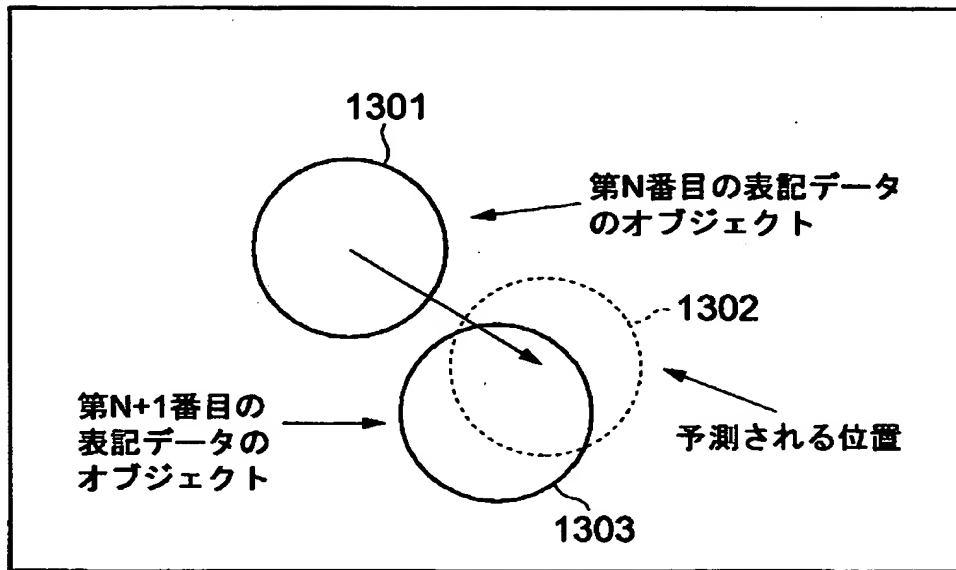
【図 11】



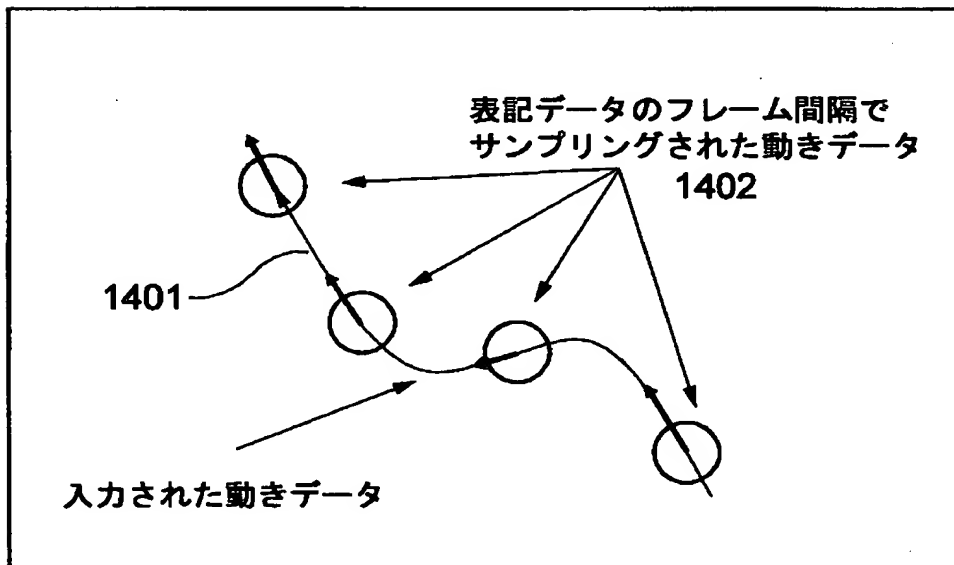
【図 1 2】



【図 13】



【図 14】



【書類名】 要約書

【要約】

【課題】 背景が動いているオブジェクトを含む映像についても検索を可能とするための映像情報記述方法を提供する。

【解決手段】 元映像からオブジェクトの位置、形状及び動きなどの情報を含む特徴量 7 0 3 と背景の動きなどの情報を含む特徴量 7 0 4 をこれらを表記データとして記述する。

【選択図】 図 7



出 願 人 履 歴 情 報

識別番号 [000003078]

|          |                  |
|----------|------------------|
| 1. 変更年月日 | 1990年 8月22日      |
| [変更理由]   | 新規登録             |
| 住 所      | 神奈川県川崎市幸区堀川町72番地 |
| 氏 名      | 株式会社東芝           |